REMARKS

Claims 46-52, 54, 55, 57, 58, and 60-71 were pending as of the action mailed on December 10, 2008. Claims 46, 49, and 51 are in independent form. The examiner withdrew claim 72 from consideration in the action mailed on December 10, 2008. This reply is being filed with a Request for Continued Examination.

Claims 47, 52, 54, 57, and 60 are being cancelled. Claims 46, 48-50, 51, 62, and 64-72 are being amended. Support for the amendments can be found in the claims and the specification, for example, in the specification at page 2, lines 2-4, page 18, lines 11-13, page 20, line 13 to page 22, line 19, page 38 line 12, and in claim 47 and 54. No new matter has been added.

Reconsideration of the action is respectfully requested in light of the foregoing amendments and the following remarks.

Restriction Requirement

The examiner required restriction between claim 72 and claims 46, 48-51, 55, and 61-71, as allegedly being combination and subcombination inventions. Because the applicant had received an action on the merits for claims 46, 48-51, 55, and 61-71, the examiner constructively elected those claims and withdrew claim 72 from consideration. The applicant traverses the restriction.

Before imposing a restriction requirement, the examiner must show that there would be a serious burden on the examiner if restriction is not required by showing that the inventions have separate classifications, that the inventions have a separate status in the art, or that the inventions require different fields of search (*see* MPEP 808.02). The applicant respectfully submits that the invention described in claims 46, 48-51, 55, and 61-71 and the invention described in claim 72 are not separately classified, do not have a separate status in the art, and do not require different fields of search.

The examiner has stated that the combination as claimed does not require the particulars of the subcombination as claimed because the examined invention does not require "identifying

query-specific duplicate documents by comparing one or more segments of the documents corresponding to the search results, wherein each of the one or more segments includes at least one of the one or more keywords," as recited in claim 72. However, the applicant submits that the recited limitation of claim 72 was also described in claims 46-48 and 54, which together describe comparing segments corresponding to search results to determine similarity between documents. Therefore, the applicant respectfully submits that the examiner has not established sufficient reasons for the restriction. In particular, the examiner has not established a serious burden in examining claim 72. The examiner has already searched and examined the features found in claim 72 through examination of claims 46-48 and 54, therefore there is no burden in examining claim 72.

Furthermore, the applicant respectfully submits that the invention described in claims 46, 48-51, 55, and 61-71 and the invention described in claim 72 are not subcombination and combination, but rather are directed to different characterizations of the same invention.

For the foregoing reasons, the applicant respectfully requests that the restriction be withdrawn. The applicant further submits that claim 72, as amended, is condition for allowance for the reasons set forth below with respect to claim 46.

Section 102 Rejections

Claims 46, 48-52, 55, 58, 61, and 64-71 were rejected under 35 U.S.C. § 102(a) as allegedly being anticipated by Zamir, Oren Eli, "Clustering Web Documents: A Phrase-Based Method for Group Search Engine Results (1999, University of Washington) ("Zamir"). The applicant respectfully disagrees.

Claim 46, as amended, recites presenting a set of final search results. This set of final search results is generated by adding a first search result to the set, determining that a first document corresponding to the first search result and a second document corresponding to a second search result are query-specific duplicate documents, and then not adding the second search result to the set of final search results.

Some of the newly added features of claim 46 were originally in claims 47 and 17. The examiner stated that Zamir disclosed these features at page 105, item 4, page 106, lines 1-9, and page 29, last paragraph.

The cited portion of Zamir on page 29 reads as follows:

> The single-pass algorithm [Hill, 68] attempt, like k means, to find spherical clusters of equal size. It is an incremental algorithm that uses a greedy search, assigning each document to a cluster only once. The first document processed is used to start the first cluster. Every additional document is compared to all existing clusters, and its most similar cluster is found. <u>If its similarity to this cluster is above a predefined threshold the document will be added to that cluster</u>, otherwise it will be used to create a new cluster. Single pass is a linear time algorithm. It suffers from being order dependent, from being extremely sensitive to the predefined threshold, and from having a tendency to promote large clusters [Rasmussen, 92]. It is, however, the most popular incremental clustering algorithm (as can be seen from its popularity in the topic detection domain [TDT, 97]). (emphasis added).

The cited portion of Zamir discloses a clustering algorithm where a document is assigned to a cluster of other documents based on its similarity to the cluster. When a document is similar to the cluster, the document is added to the cluster. In contrast, claim 46 recites <u>not adding</u> a second search result to the set of final search results after determination that a document corresponding to the second search result and a document corresponding to a search result already in the set of final search results are query-specific duplicate documents.

The cited portions of Zamir on pages 105 and 106 state:

> When analyzing Web search results, large frequent sets are often found when a certain document appears several times in slightly different variations. If the document is of considerable length, this will lead to a very expensive computation. We therefore employ a "near duplicate elimination" scheme in which we try to identify documents which are highly similar, then remove one of them from consideration and simply add it to the cluster that its "duplicate" ended up in. This can be done before the frequency set discovery phase, but that would be quadratic in the number of documents. We employ an alternatively [sic] technique which turns out to be much faster in practice. When too many (over 10000) frequent sets are discovered in a certain step of the algorithm (at a 4% support level), we perform the following duplicate identification procedure. We make a pass through all the frequent sets and calculate for each document pair its frequent sets intersection ratio—the number of frequent sets that the two documents appear in together out of the total number of frequent sets that either one of then [sic] appears in. If this ratio is large, we mark these two documents as duplicate and remove one of them from consideration. This

Applicant : Benedict Anthony Gomes   Attorney's Docket No.: 16113-1571001
Serial No. : 10/602,965
Filed   : June 24, 2003
Page   : 12 of 14

process is linear in the number of frequent sets at that stage, but quadratic in the length of the longest document list of a frequent set. This is an expensive process, but it is much faster in practice, for large document sets, than finding near-duplicates by comparing all pairs of documents to each other.

The cited portions of Zamir disclose an optimization to a clustering algorithm where documents that are duplicates of a given document are temporarily ignored during clustering, and then added back into to the cluster corresponding to the given duplicate before the results are presented to the user. The optimization taught in Zamir is a temporary step and does not affect the results presented to the user (i.e., the final clusters include duplicates).

Furthermore, Zamir teaches away from a presenting filtered search results that do not include identified duplicates. The objective in Zamir is the opposite: to provide clusters of similar documents including duplicates. Consequently, the relied upon temporary removal of duplicate documents is simply for computational efficiency and not to remove duplicates form the final clusters that are presented.

Thus, Zamir does not disclose or suggest adding a first search result presenting a set of final search results generated by adding a first search result to the set of final search results, determining that a first document corresponding to the first search result and a second document corresponding to the second search result are query-specific duplicate documents, and not adding the second search result to the set of final search results, as recited in claim 46.

The applicant submits that claim 46, as well as claims 48, 55, 62, and 64-67, which depend from claim 46, are in condition for allowance for at least the foregoing reasons.

Claim 49 includes features corresponding to those of claim 46 and was rejected for the same reasons. Therefore, the applicant submits that claim 49, as well as claims 50 and 58, which depends from claim 49, are allowable for at least the same reasons set forth above with respect to claim 46.

Claim 51, as amended, recites a final results presenter for presenting a set of final search results and a final results generator for generating the set of final search results. The final result generator generates the set of final search results by adding a first search result to the set of final search results, determining whether a first document corresponding to the first search result and a second document corresponding to a second search result are query-specific duplicate

documents, and adding the second search result to the set of final search results when the similarity determination facility determines that the first document and the second document are not query-specific duplicate documents and not adding the second search result to the set of final search results when the similarity determination facility determines the first document and the second document are query-specific duplicate documents.

Claim 51 was rejected for the same reasons claim 46 was rejected. The applicant respectfully disagrees with the rejection. As described above in reference to claim 46, Zamir discloses clustering documents by adding a document to a set containing similar documents. In contrast, claim 51 recites adding a search result to the set of final search results when it is not a query-specific duplicate of another document and not adding a search result to the set of final search results when it is not a query-specific duplicate of another document. While Zamir does disclose a temporary filtering step, the filtering taught by Zamir does not affect the results presented to the user. Therefore, Zamir does not disclose or suggest the features of claim 51.

The applicant submits that claim 51, as well as claims 52, 61, 63, and 68-71, which depend from claim 51, are in condition for allowance for at least the foregoing reasons.

Conclusion

For the foregoing reasons, the applicant submits that all the claims are in condition for allowance.

By responding in the foregoing remarks only to particular positions taken by the examiner, the applicant does not acquiesce with other positions that have not been explicitly addressed. In addition, the applicant's selecting some particular arguments for the patentability of a claim should not be understood as implying that no other reasons for the patentability of that claim exist. Finally, the applicant's decision to amend or cancel any claim should not be understood as implying that the applicant agrees with any positions taken by the examiner with respect to that claim or other claims.

Please apply any charges or credits to deposit account 06-1050.

Respectfully submitted,

Date: 3/10/09

Elspeth S. White
Reg. No. 62,133

Fish & Richardson P.C.
500 Arguello Street, Suite 500
Redwood City, California 94063
Telephone: (650) 839-5070
Facsimile: (877) 769-7945

50639635.doc